Sequential Beat-to-Beat P and T Wave Delineation and Waveform Estimation in ECG Signals: Block Gibbs Sampler and Marginalized Particle Filter

Chao Lin*, Georg Kail†, Audrey Giremus‡, Corinne Mailhes§1,4, Jean-Yves Tourneret¶4, and Franz Hlawatsch∥2

1TeSA lab, Toulouse, France
2Institute of Telecommunications, Vienna University of Technology, Vienna, Austria
3IMS-LAPS-UMR 5218, University of Bordeaux, Talence, France
4IRIT Lab-UMR 5505, University of Toulouse, Toulouse, France

October 27, 2013

Abstract

For ECG interpretation, the detection and delineation of P and T waves are challenging tasks. This report proposes Bayesian methods for simultaneous detection, threshold-free delineation, and waveform estimation of P and T waves on a beat-to-beat basis. By contrast to state-of-the-art methods that process multiple-beat signal blocks, the proposed Bayesian methods account for beat-to-beat waveform variations by sequentially estimating the waveforms for each beat. Our methods are based on a Bayesian signal model that takes into account previous beats as prior information. To estimate the unknown parameters of the Bayesian model, we first propose a block Gibbs sampler that exhibits fast convergence in spite of the strong local dependencies in the ECG signal. Then, in order to take into account all the information contained in the past rather than considering only one previous beat, a sequential Monte Carlo method is studied, with a marginalized particle filter that efficiently estimates the unknown parameters of the dynamic model. Both methods are evaluated on the annotated QT database. Performance comparisons demonstrate that both methods achieve significant improvements in detection rate and delineation accuracy compared to state-of-the-art methods, thus providing promising approaches for sequential P and T wave analysis problems in ECG signals.

1 Context of the work

This work is a joint work between several signal processing research labs: TeSA lab (Toulouse, France), Institute of Communications (Vienna, Austria) and two French CNRS labs IMS and IRIT (Bordeaux and Toulouse, France). It has been supported by a grant of Saint Jude Med. for TeSA Lab, by the Austrian Science Fund (FWF) under grant S10603 for the Institute of

*chao.lin@tesa.prd.fr
†georg.kail@nt.tuwien.ac.at
‡audrey.giremus@ims-bordeaux.fr
§corinne.mailhes@tesa.prd.fr
¶Jean-Yves.Tourneret@enseeiht.fr
∥franz.hlawatsch@tuwien.ac.at
Telecommunications and made within the project Dynbrain funded by STIC AmSud for the French CNRS IRIT.

2 Introduction

The electrocardiogram (ECG) represents the electrical activity of the heart which corresponds to repetitions of a cardiac cycle, i.e., a heartbeat. Each heartbeat consists of a QRS complex surrounded by P and T waves that are associated with the mechanical phases occurring during a cardiac cycle. Most of the clinically useful information can be derived from the wave intervals, amplitudes, and morphology. Therefore, the development of efficient and robust methods for automatic ECG delineation (determining the locations of the peaks and boundaries of the individual waves) has become a major challenge for the biomedical signal processing community.

Among the ECG waves, the QRS complex is relatively easy to detect and is thus generally used as a reference within the cardiac cycle. For P and T wave detection and delineation, most algorithms perform QRS detection first and then define temporal search windows before and after the QRS location points in which they assume the P and T waves. Subsequently, an appropriate strategy is used to enhance the distinctive features of each wave in order to locate the wave peaks and boundaries.

In the last two decades, a variety of techniques have been proposed for automatically detecting and delineating P and T waves [1–8]. These techniques are based on adaptive filtering [29], low-pass differentiation [16], wavelet transform [18, 22], action potential models [32], pattern recognition [31], extended Kalman filters [26], or evolutionary optimization [10]. However, because of the low slope and amplitude of the P and T waves as well as the presence of noise, interference, and baseline fluctuation, P and T wave detection and delineation remain challenging tasks. Furthermore, in addition to the locations of the wave peaks and boundaries, the shapes and amplitudes of P and T waves have also been shown to contain important information about numerous pathologies [5].

A Bayesian model was recently proposed to simultaneously solve the P and T wave delineation and waveform estimation problems [20, 21]. This model takes into account prior distributions for the unknown parameters (wave locations and amplitudes [20] as well as waveform and local baseline coefficients [21]). These prior distributions were combined with the likelihood of the observed data to obtain the posterior distribution of the unknown parameters. Several Gibbs-type samplers were then proposed to cope with the complexity of this posterior distribution and estimate the model parameters [20, 21]. However, the Bayesian model of [20, 21] relied on a non-overlapped multiple-beat processing window. More precisely, the shapes of the P and T waves within a multiple-beat processing window were assumed to be equal, whereas their amplitudes and locations were allowed to vary from one beat to another. Due to the pseudo-cyclostationary nature of the ECG signal, the P and T waveforms in a given beat are usually similar but not exactly equal to those of the adjacent beats. Therefore, the performance of P and T wave delineation can be expected to improve if the waveforms are estimated in a beat-to-beat manner that allows for temporal variations of waveform morphology across the beats. A beat-to-beat processing mode is also advantageous for an on-line operation with reduced memory requirements and rapid adaptation to changing signal characteristics.

In this report, we present and study Bayesian methods that enable simultaneous P and T wave delineation and waveform estimation on a beat-to-beat basis. First, a beat-to-beat Bayesian model is proposed which modifies the multiple-beat-window-based model studied in [20, 21] by introducing dependencies among waveform coefficients. Instead of assigning a white Gaussian prior to the temporal sequence of waveform coefficients, we use a prior “with memory” that depends on the estimates of the previous beat. A Gibbs sampler with a block constraint,
referred to as block Gibbs sampler (BGS), is then used for estimating the parameters of the resulting beat-to-beat model. Simulation results show that the proposed sequential model and processing improve the convergence behavior of the samplers proposed in [20, 21] as well as the accuracy of estimating the locations, amplitudes, and shapes of the P and T waves. The improved convergence behavior can be explained by a considerable reduction of the parameter dimension, since only one beat is processed at any time instant instead of multiple beats.

In the second part of this report, we present a sequential Monte Carlo method that takes into account all the information contained in the past rather than only that of the previous beat. The principle of this method is to exploit the sequential nature of the ECG signal by defining an appropriate dynamic model. This model adapts to the morphology variations across the ECG beats by using a random walk model for the waveform coefficients. A particle filter is then employed to estimate the unknown parameters of the proposed model. The key idea of the particle filter is to represent the posterior density of interest by a set of random samples with associated weights and to compute parameter estimates from these samples and weights. Despite the simplicity of the particle filter principle, its main drawback is its computational complexity, especially for a large state dimension. In practice, if the state dimension is high, a lot of random samples are necessary to achieve a good accuracy of the estimates. However, this problem can be alleviated for nonlinear models containing a subset of parameters which are linear and Gaussian, conditional upon the other parameters. In this case, the linear/Gaussian parameters can be optimally estimated through standard linear Gaussian filtering. This technique is often referred to as Rao-Blackwellization [8] or marginalization [27]. In our case, the state equations are linear with respect to a subset of the unknown parameters. Thus, we propose to use a marginalized particle filter (MPF) that generates particles in the space of the “nonlinear” parameters and runs one Kalman filter for each of these particles to estimate the “linear” parameters.

A comparison between the proposed sequential BGS, the proposed MPF, and state-of-the-art methods shows that both of the proposed methods provide significant improvements in terms of estimation performance for the locations, amplitudes, and shapes of the P and T waves, with the MPF method typically exhibiting a slightly better performance than the BGS method at the cost of a higher complexity.

The report is organized as follows. Section 3 describes the proposed beat-to-beat Bayesian model for the non-QRS signal components, a BGS that generates samples distributed according to the posterior of this Bayesian model, and the detectors and estimators based on the generated samples. Section 4 presents a dynamic model based on the proposed beat-to-beat Bayesian framework and an associated MPF. Section 5 reports the results of numerical simulations performed on the standard annotated QT database [17]. These results allow the performance of the two proposed methods to be compared with that of state-of-the-art algorithms. Finally, Section 6 presents conclusions and suggests future work.

3 Beat-to-beat Bayesian model and block Gibbs sampler

3.1 Signal model for one non-QRS interval

It is common to partition ECGs into QRS complexes and non-QRS intervals. Non-QRS intervals are located between the end of a QRS complex and the subsequent QRS onset, and they potentially contain P and T waves. In this report, we assume that the locations of the non-QRS intervals have been determined by a preliminary QRS detection step, using, e.g., the Pan-Tompkins algorithm [23]. We also assume that any baseline wanderings have been removed by, e.g., the median filtering technique proposed in [4].

As shown in Fig. 1, the non-QRS interval \( J_n \) associated with the \( n \)th heartbeat consists
of two complementary subintervals: a T search interval $\mathcal{J}_{T,n}$, which may contain a T wave, and a P search interval $\mathcal{J}_{P,n}$, which may contain a P wave. The temporal lengths of the intervals $\mathcal{J}_n$, $\mathcal{J}_{T,n}$, and $\mathcal{J}_{P,n}$ will be denoted by $N_n$, $N_{T,n}$, and $N_{P,n}$, respectively. Note that $N_{T,n} + N_{P,n} = N_n$. The lengths $N_{T,n}$ and $N_{P,n}$ can be determined by a cardiologist or simply as fixed percentages of $N_n$. In this work, we choose $N_{T,n} = N_{P,n} = N_n/2$. Our goal is to estimate the locations, amplitudes, and shapes of the P and T waves within their respective search intervals $\mathcal{J}_{T,n}$ and $\mathcal{J}_{P,n}$. Note that only the locations of the wave peaks are constrained to lie within their respective search intervals.

### 3.1.1 Convolution model

The baseline-free signal in the non-QRS interval $\mathcal{J}_n$ can be approximated by two pulses representing the P and T waves (see Fig. 1). Similar to the blind deconvolution problem in [3, 14], the T wave is modeled by the convolution of an unknown binary “indicator sequence” $b_{T,n} = (b_{T,n,1}, \ldots, b_{T,n,N_{T,n}})^T$ indicating the wave locations ($b_{T,n,k} = 1$ if there is a wave at the $k$th possible location, $b_{T,n,k} = 0$ otherwise) with an unknown T waveform $h_{T,n} = (h_{T,n,-L}, \ldots, h_{T,n,L})^T$. Analogous definitions for the P wave yield $b_{P,n} = (b_{P,n,1}, \ldots, b_{P,n,N_{P,n}})^T$ and $h_{P,n} = (h_{P,n,-L}, \ldots, h_{P,n,L})^T$. Here, the waveform length $2L + 1$ is chosen as a fixed percentage of $N_n$ that is large enough to accommodate the actual supports of the P and T waves. Within each indicator vector $b_{T,n}$ and $b_{P,n}$, at most one entry is nonzero because at most one wave may occur in any given search interval. According to this model, the $n$th non-QRS signal component can be expressed as follows:

$$x_{n,k} = \sum_{j=1}^{N_{T,n}} h_{T,n,k-j} b_{T,n,j} + \sum_{j=N_{T,n}+1}^{N_n} h_{P,n,k-j} b_{P,n,j-N_{T,n}} + w_{n,k}, \quad k \in \mathcal{J}_n = \{1, \ldots, N_n\}. \quad (1)$$

Here, $w_{n,k}$ denotes white Gaussian noise with unknown variance $\sigma^2_{w,n}$. Furthermore, we have set $h_{T,n,k} = h_{P,n,k} = 0$ for $k \notin \{-L, \ldots, L\}$.

![Figure 1: Signal model for the beat-to-beat processing scheme.](image-url)
3.1.2 Waveform expansion

Following [28, 13], we represent the P and T waveforms by a basis expansion using discrete-time versions of Hermite functions. Thus, the waveform vectors can be written as

$$h_{T,n} = H_{T,n} \alpha_{T,n}, \quad h_{P,n} = H_{P,n} \alpha_{P,n}$$

where $H$ is a $(2L + 1) \times G$ matrix whose columns are the first $G$ Hermite functions (with $G \leq 2L+1$), suitably sampled and truncated to length $2L+1$, and $\alpha_{T,n}$ and $\alpha_{P,n}$ are unknown coefficient vectors of length $G$. By using these expansions, the number of unknown parameters can be significantly reduced (from $2L+1$ to $G$ for each waveform). Note that the amplitudes of the P and T waves are absorbed into the coefficient vectors $\alpha_{T,n}$ and $\alpha_{P,n}$. This is a difference from the model in [20, 21], where the amplitudes were defined for each heartbeat individually whereas the P and T waveforms were fixed for multiple heartbeats.

3.1.3 Vector formulation

Using (2), we obtain the following vector representation of the non-QRS signal in (1):

$$x_n = B_{T,n} H_{T,n} \alpha_{T,n} + B_{P,n} H_{P,n} \alpha_{P,n} + w_n$$

where $x_n = (x_{n,1} \cdots x_{n,N_n})^T$, $B_{T,n}$ is the $N_n \times (2L + 1)$ Toeplitz matrix with first row and column respectively $(b_{T,n,L+1} \cdots b_{T,n,1} 0 \cdots 0)$ and $(b_{T,n,L+1} \cdots b_{T,n,N_n} 0 \cdots 0)^T$, $B_{P,n}$ is the $N_n \times (2L + 1)$ Toeplitz matrix with last row $(0 \cdots 0 b_{P,n,N_p,1} \cdots b_{P,n,N_p,-L})^T$ and last column $(0 \cdots 0 b_{P,n,1} \cdots b_{P,n,N_p,N_p-L})^T$, and $w_n = (w_{n,1} \cdots w_{n,N_n})^T$ is a Gaussian vector with zero mean and covariance matrix $\sigma_{w,n}^2 I_{N_n}$, with $I_{N_n}$ denoting the identity matrix of size $N_n \times N_n$.

3.2 Likelihood function, prior, and posterior

According to the parametrization introduced in Section 3.1, the unknown parameters for the $n$th non-QRS interval $\mathcal{J}_n$ are given by the random vector $\theta_n = (b_{T,n}^T b_{P,n}^T \alpha_{T,n}^T \alpha_{P,n}^T \sigma_{w,n}^2)^T$. Note, in particular, that the noise variance $\sigma_{w,n}^2$ may vary from one heartbeat to another. Bayesian detection/estimation relies on the posterior distribution

$$p(\theta_n|x_n) \propto p(x_n|\theta_n)p(\theta_n)$$

where $\propto$ means “equal up to a positive factor that does not depend on $\theta_n$,” $p(x_n|\theta_n)$ is the likelihood function, and $p(\theta_n)$ is the prior distribution of $\theta_n$. The next two subsections present the likelihood function and priors considered in this study.

3.2.1 Likelihood function

Using (3) and the fact that $w_{n,k}$ is white and Gaussian with variance $\sigma_{w,n}^2$, the likelihood function (viewed as a function of $x_n$) is obtained as

$$p(x_n|\theta_n) = \mathcal{N}(B_{T,n} H_{T,n} \alpha_{T,n} + B_{P,n} H_{P,n} \alpha_{P,n}, \sigma_{w,n}^2 I_{N_n})$$

where $\mathcal{N}(\mu, C)$ denotes the multivariate Gaussian probability density function with mean vector $\mu$ and covariance matrix $C$. 

5
3.2.2 Prior distributions

Wave indicators The indicators \( b_{T,n,k} \) are subject to a block constraint: within \( J_{T,n} \), there is one T wave (thus \( \|b_{T,n}\| = 1 \)) or none (thus, \( \|b_{T,n}\| = 0 \)), the latter case being very unlikely. Therefore, we define the prior of \( b_{T,n} \) as

\[
p(b_{T,n}) = \begin{cases} 
  p_0 & \text{if } \|b_{T,n}\| = 0 \\
  p_1 & \text{if } \|b_{T,n}\| = 1 \\
  0 & \text{otherwise}
\end{cases}
\]  

where \( p_1 = (1 - p_0)/N_{T,n} \) and \( p_0 \) is chosen very small. Note that there are \( N_{T,n} \) vectors satisfying \( \|b_{T,n}\| = 1 \) whereas the zero vector is the only vector satisfying \( \|b_{T,n}\| = 0 \). Thus, the probabilities in (5) sum to one. Similarly, within \( J_{P,n} \), there is one P wave or none; therefore, the prior of \( b_{P,n} \) is defined as in (5), with \( p_1 = (1 - p_0)/N_{P,n} \). The wave indicator vectors \( b_{T,n} \) and \( b_{P,n} \) for different search intervals (i.e., different values of \( n \)) are assumed to be statistically independent.

Waveform coefficients The waveform coefficient vectors \( \alpha_{T,n} \) and \( \alpha_{P,n} \) for the \( n \)th non-QRS interval \( J_n \) are supposed to depend on the respective coefficient vectors in the \((n-1)\)th non-QRS interval \( J_{n-1} \). Consider the T wave as an example. The prior of \( \alpha_{T,n} \) is defined as

\[
p(\alpha_{T,n}|b_{T,n}, \alpha_{T,n-1}) = \begin{cases} 
  \delta(\alpha_{T,n} - \alpha_{T,n-1}) & \text{if } \|b_{T,n}\| = 0 \\
  \mathcal{N}(\alpha_{T,n-1}, \sigma_\alpha^2 I_G) & \text{if } \|b_{T,n}\| = 1
\end{cases}
\]

where \( \delta(\cdot) \) is the Dirac delta function. For the variance \( \sigma^2_\alpha \), we choose a value that yields a reasonable variability of the waveform coefficients from one interval to another. Note that when there is no T wave in the search interval (\( \|b_{T,n}\| = 0 \)), the prior sets \( \alpha_{T,n} \) equal to \( \alpha_{T,n-1} \), i.e., the waveform coefficients are equal to those in the previous interval \( J_{n-1} \). The prior of the P waveform coefficient vector \( \alpha_{P,n} \) is defined in an analogous way, with \( \alpha_{T,n-1} \) replaced by \( \alpha_{P,n-1} \). These definitions of the priors of \( \alpha_{T,n} \) and \( \alpha_{P,n} \) introduce a memory in the statistical model for the P and T waveforms and, in turn, induce a sequential processing.

Noise variances The noise variances \( \sigma^2_{w,n} \) are modeled as independent random variables distributed according to inverse gamma distributions \( p(\sigma^2_{w,n}) = IG(\xi, \eta) \), where \( \xi \) and \( \eta \) are fixed hyperparameters defining a vague prior (as in [6]).

We note at this point that the Gaussian priors of \( \alpha_{T,n} \) and \( \alpha_{P,n} \) are conjugate priors with respect to the Gaussian likelihood function (4), i.e., the resulting full conditional distributions (required in the Gibbs sampler) are also Gaussian [24, p. 97]. A similar remark applies to the inverse gamma prior of \( \sigma^2_{w,n} \). The choice of conjugate priors yields a considerable simplification of our detection/estimation algorithm.

Joint prior Since there are no known relations between \( (b_{T,n}, \alpha_{T,n}), (b_{P,n}, \alpha_{P,n}) \), and \( \sigma^2_{w,n} \), all these sets of parameters are assumed to be a priori statistically independent. Therefore, the joint prior for the total parameter vector \( \theta_n = (b_{T,n}^T, b_{P,n}^T, \alpha_{T,n}^T, \alpha_{P,n}^T, \sigma^2_{w,n})^T \) factors as

\[
p(\theta_n|\alpha_{T,n-1}, \alpha_{P,n-1}) = p(\alpha_{T,n}|b_{T,n}, \alpha_{T,n-1}) p(b_{T,n}) p(\alpha_{P,n}|b_{P,n}, \alpha_{P,n-1}) p(b_{P,n}) p(\sigma^2_{w,n}).
\]  

3.2.3 Posterior distribution

The posterior distribution of the parameter vector \( \theta_n \) is obtained by using Bayes’ rule, i.e.,

\[
p(\theta_n|x_n, \alpha_{T,n-1}, \alpha_{P,n-1}) \propto p(x_n|\theta_n) p(\theta_n|\alpha_{T,n-1}, \alpha_{P,n-1})
\]

\(^1\| \cdot \| \) denotes the \( \ell_2 \) norm, i.e., \( \|x\|^2 = x^T x \).
where the right hand term can be further expressed and factored using (4) and (7). Because the proposed method works sequentially and all estimates from the previous beat are available, we can substitute the estimates \( \hat{\alpha}_{T,n-1} \) and \( \hat{\alpha}_{p,n-1} \) for \( \alpha_{T,n-1} \) and \( \alpha_{p,n-1} \) in \( p(\theta_n|\mathbf{x}_n, \alpha_{T,n-1}, \alpha_{p,n-1}) \) when estimating \( \theta_n \) based on (8).

Due to the complexity of the posterior distribution, we propose to use a Monte Carlo (sample-based) detection/estimation method. More specifically, we propose a BGS that generates samples asymptotically distributed according to \( p(\theta_n|\mathbf{x}_n, \hat{\alpha}_{T,n-1}, \hat{\alpha}_{p,n-1}) \) (see Section 3.3). From these samples, the discrete parameters \( \mathbf{b}_{T,n} \) and \( \mathbf{b}_{p,n} \) are then detected by means of the sample-based maximum a posteriori (MAP) detector, and the continuous parameters \( \alpha_{T,n}, \alpha_{p,n}, \) and \( \sigma^2_{w,n} \) are estimated by means of the sample-based minimum mean square error (MMSE) estimator, as described in Section 3.4.

3.3 Block Gibbs sampler for beat-to-beat wave extraction

3.3.1 Block Gibbs sampler algorithm

The proposed BGS for the \( n \)th non-QRS interval \( J_n \) is summarized in Algorithm 1. Note that, in the rest of Section 3, the interval index \( n \) is omitted for all parameters to simplify the notation, while the index \( n - 1 \) is kept to avoid any ambiguity.

**Algorithm 1** Block Gibbs sampler

Sample \( \mathbf{b}_T \) from \( p(\mathbf{b}_T|\mathbf{b}_P, \hat{\alpha}_{T,n-1}, \alpha_{P}, \sigma_w^2, \mathbf{x}) \)

Sample \( \alpha_T \) from \( p(\alpha_T|\mathbf{b}_T, \mathbf{b}_P, \hat{\alpha}_{T,n-1}, \alpha_{P}, \sigma_w^2, \mathbf{x}) \)

Sample \( \mathbf{b}_P \) from \( p(\mathbf{b}_P|\mathbf{b}_T, \hat{\alpha}_{P,n-1}, \alpha_{T}, \sigma_w^2, \mathbf{x}) \)

Sample \( \alpha_P \) from \( p(\alpha_P|\mathbf{b}_T, \mathbf{b}_P, \hat{\alpha}_{P,n-1}, \alpha_{T}, \sigma_w^2, \mathbf{x}) \)

Sample \( \sigma_w^2 \) from \( p(\sigma_w^2|\mathbf{b}_T, \mathbf{b}_P, \alpha_T, \alpha_P, \mathbf{x}) \)

The term “block Gibbs sampler” is used to reflect the block constraints related to the wave indicator vectors \( \mathbf{b}_T \) and \( \mathbf{b}_P \), which are encompassed in the corresponding priors (see (5)). To see that Algorithm 1 is a valid Gibbs sampler, note that the sampling steps for \( \mathbf{b}_T \) and \( \alpha_T \) are equivalent to jointly sampling \( \mathbf{b}_T \) and \( \alpha_T \) from \( p(\mathbf{b}_T, \alpha_T|\mathbf{b}_P, \hat{\alpha}_{T,n-1}, \alpha_{P}, \sigma_w^2, \mathbf{x}) \), and similarly for \( \mathbf{b}_P \) and \( \alpha_P \). The sampling distributions used in Algorithm 1 are specified in the following section; their derivations are provided in Appendix A.

3.3.2 Sampling distributions

**Wave indicators** The sampling distribution for the T wave indicator vector \( \mathbf{b}_T \) is

\[
p(\mathbf{b}_T|\mathbf{b}_P, \hat{\alpha}_{T,n-1}, \alpha_{P}, \sigma_w^2, \mathbf{x}) \propto \begin{cases} p_0 & \text{if } \|\mathbf{b}_T\| = 0 \\ \sigma_{\alpha}^{-G} \sqrt{\Sigma_1} \exp(\mu_1 \Sigma_1^{-1} \mu_1) p_1 & \text{if } \|\mathbf{b}_T\| = 1 \\ 0 & \text{otherwise} \end{cases}
\]

with

\[
\mu_1 = \Sigma_1 \left( \frac{H^T B^T \tilde{x}_T}{\sigma_w^2} + \frac{\hat{\alpha}_{T,n-1}}{\sigma_{\alpha}^2} \right)
\]

\[
\Sigma_1 = \left( \frac{H^T B^T B H}{\sigma_w^2} + \frac{I_G}{\sigma_{\alpha}^2} \right)^{-1}
\]
Here, $\hat{x}_T = x - B_T H \alpha_T$. The sampler evaluates all hypotheses for $b_T$ conditioned on the current samples of all other parameters. There are $N_T + 1$ such hypotheses, because, according to (9), $b_T$ has either no 1-entry or exactly one 1-entry at one of $N_T$ possible locations within $J_T$.

The sampling distribution for the P wave indicator vector, $p(b_T, b_P, \hat{\alpha}_{T,n-1}, \alpha_P, \sigma_w^2, x)$, is obtained in an analogous manner, with $\hat{\alpha}_{T,n-1}$ replaced by $\hat{\alpha}_P,n-1$, $B_T$ replaced by $B_P$, and $\hat{x}_T$ replaced by $\hat{x}_P = x - B_T H \alpha_T$.

**Waveform coefficients** The sampling distribution for $\alpha_T$ is

$$p(\alpha_T | b_T, b_P, \hat{\alpha}_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \begin{cases} \delta(\alpha_T - \hat{\alpha}_{T,n-1}) & \text{if } ||b_T|| = 0, \\ N(\mu_1, \Sigma_1) & \text{if } ||b_T|| = 1 \end{cases}$$

with $\mu_1$ and $\Sigma_1$ as defined above. In particular, the samples of $\alpha_T$ are given by $\hat{\alpha}_{T,n-1}$ if there is no T wave in the search interval ($||b_T|| = 0$). The sampling distribution $p(\alpha_P | b_T, b_P, \hat{\alpha}_{P,n-1}, \alpha_T, \sigma_w^2, x)$ is obtained similarly for $\alpha_P$.

**Noise variance** The sampling distribution for $\sigma_w^2$ is well known (e.g., [15]) to be the following inverse gamma distribution:

$$p(\sigma_w^2 | b_T, b_P, \alpha_T, \alpha_P, x) = IG(\xi', \eta')$$

with

$$\xi' = \xi + \frac{N}{2}, \quad \eta' = \eta + \frac{1}{2} ||x - B_T H \alpha_T - B_P H \alpha_P||^2.$$
indices $i \in \{1, \ldots, N_s\}$ such that both $b_T^{(i)} = \hat{b}_T$ and $b_P^{(i)} = \hat{b}_P$. To estimate $\theta_{\sim b}$, we use the sample mean

$$\hat{\theta}_{\sim b} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \theta_{\sim b}^{(i)}$$

where $\theta_{\sim b}^{(i)} \triangleq \left( \alpha_{T}^{(i)T} \alpha_{P}^{(i)} \sigma_w^{2(i)} \right)^T$ and $|\mathcal{I}|$ denotes the number of elements in $\mathcal{I}$. This can be interpreted as a sample-based approximation of the MMSE estimator (note that the MMSE estimator is given by the posterior mean $E\{\theta_{\sim b} \mid x, b_T, b_P, \hat{\alpha}_T, n^{\prime}, \hat{\alpha}_P, n^{\prime}\}$). Thus, $\hat{\theta}_{\sim b}$ depends on $b_T, b_P, \hat{\alpha}_{T,n^{\prime}}$, and $\hat{\alpha}_{P,n^{\prime}}$.

### 3.4.3 Wave delineation

Because of the convolution model (1) and the fact that the P and T waveforms are unknown, our detection/estimation problem is essentially a blind deconvolution problem, and as such it is affected by a time-shift ambiguity [3, 15]. In fact, the detection/estimation results obtained for the wave indicators $b_T,k$ and $b_P,k$ and the waveforms $h_T,k$ and $h_P,k$ (as represented by the waveform coefficients $\alpha_T$ and $\alpha_P$) are ambiguous with respect to their relative locations, in the sense that a temporal shift of $b_T,k$ and $b_P,k$ can be compensated for by an inverse temporal shift of $h_T,k$ and $h_P,k$.

For wave delineation (localization of the peaks and boundaries of the P and T waves), it is necessary to resolve this time-shift ambiguity. Following [3], this can be achieved by performing an appropriate time shift after generating the waveform samples in the block Gibbs sampler. This time shift ensures that the maximum of the waveform is located at the center $k=0$ of the waveform support interval $\{-L, \ldots, L\}$ and, thus, the location of a nonzero detected indicator $\hat{b}_T,k = 1$ or $\hat{b}_P,k = 1$ directly indicates the peak of the respective T or P wave. Note that a detailed description of an algorithm for resolving the time-shift ambiguity is also provided in [15].

The wave delineation consists of determining the boundaries of the detected P and T waves. It is broadly accepted that the turning points defined by the largest local maximum of the curvature of the estimated waveform on each side of the detected wave peak are good estimates of the wave boundaries [31, 2]. The curvature of, e.g., the estimated T waveform $\hat{h}_{T,k}$ is defined as [31]

$$\kappa_{T,k} \triangleq \frac{\hat{h}_{T,k}''}{\left[1 + (\hat{h}_{T,k}')^2\right]^{3/2}}, \quad k \in \{-L, \ldots, L\}$$

where $\hat{h}_{T,k}'$ and $\hat{h}_{T,k}''$ are discrete-time counterparts of the first and second derivatives (e.g., $\hat{h}_{T,k}'$ is defined as the difference $\hat{h}_{T,k} - \hat{h}_{T,k-1}$).

The delineation method described above avoids the use of rigid detection and delineation thresholds. Fig. 2 illustrates the method by showing the delineation results obtained for three different T wave morphologies.

Simulation results for the proposed BGS will be presented in Section 5.

### 4 Marginalized particle filter

Section 3 presented a beat-to-beat Bayesian model that introduced dependencies among waveform coefficients. A prior “with memory” (depending on the previous estimates of the P and T waveforms) was assigned to the current beat. In this section, elaborating on [19], an MPF method [7] is proposed to take into account all the information contained in the past of the
current beat to be processed. First, we present a dynamic model as a basis for performing simultaneously P and T wave delineation and waveform estimation on a beat-to-beat basis. This dynamic model is similar to the Bayesian model introduced in Section 3. However, it adapts to the morphology variations across the ECG beats by using a random walk model for the waveform coefficients. Then, following the sequential Monte Carlo principle, an MPF is introduced to estimate the unknown parameters of the proposed model. The key idea is to generate particles only for the states appearing non-linearly in the dynamics and run one Kalman filter for each of these particles to estimate the “linear” parameters.

4.1 Dynamic signal model for non-QRS intervals

As in Section 3, we assume that the locations of the non-QRS intervals have been determined by a preliminary QRS detection step and that baseline wanderings have been removed by another preprocessing stage.

The signal model is the same as in Section 3.1, except for the following two differences. First, the P and T waves are cut off at the edges of the respective intervals \( J_{T,n} \) and \( J_{P,n} \), which ensures that each wave is entirely contained within its interval. This leads to the following simplification of (1):

\[
x_{n,k} = \sum_{j=1}^{N_{T,n}} h_{T,n,k-j} b_{T,n,j} + w_{n,k}, \quad k \in J_{T,n} = \{1, \ldots, N_{T,n}\}
\]

\[
x_{n,k} = \sum_{j=1}^{N_{P,n}} h_{P,n,k-j-N_{T,n}} b_{P,n,j} + w_{n,k}, \quad k \in J_{P,n} = \{N_{T,n}+1, \ldots, N_n\}.
\]

Defining \( \mathbf{x}_{T,n} = (x_{n,1} \cdots x_{n,N_{T,n}})^T \) and using (2), we obtain the following vector representation of the T wave interval in (13):

\[
\mathbf{x}_{T,n} = \tilde{\mathbf{B}}_{T,n} \mathbf{H} \alpha_{T,n} + \mathbf{w}_{T,n}
\]

where \( \tilde{\mathbf{B}}_{T,n} \) comprises the first \( N_{T,n} \) rows of \( \mathbf{B}_{T,n} \). A similar vector representation can be obtained for the signal vector of the P wave interval, \( \mathbf{x}_{P,n} = (x_{n,N_{T,n}+1} \cdots x_{n,N_n})^T \) using \( \tilde{\mathbf{B}}_{P,n} \), which comprises the last \( N_{P,n} \) rows of \( \mathbf{B}_{P,n} \).
Second, the variance of the noise $w_{n,k}$ is now assumed to be constant over all heartbeats ($\sigma^2_{w,n} \equiv \sigma^2_w$) and known. Thus, $w_{T,n} = (w_{n,1} \cdots w_{n,N_T,n})^T$ and $w_{P,n} = (w_{n,N_T,n+1} \cdots w_{n,N_P,n})^T$ are Gaussian vectors with zero mean and covariance matrix $\sigma^2_w I_{N_T,n}$ and $\sigma^2_w I_{N_P,n}$, respectively.

4.2 Likelihood function, posterior, and prior

Using the modified signal model from Section 4.1, the likelihood function—now taking into account all time instants up to $n$—factors as

$$p(x_{1:n} | \theta_{0:n}) = p(x_{T,1:n} | b_{T,0:n}, \alpha_{T,0:n}) p(x_{P,1:n} | b_{P,0:n}, \alpha_{P,0:n}).$$

(15)

Here, e.g., $x_{1:n} \triangleq (x_1^T \cdots x_n^T)^T$ and $\theta_{0:n} \triangleq (\theta_{0}^T \cdots \theta_{n}^T)^T$. As before (cf. Section 3.2.2), we assume that the T wave parameters are independent of the P wave parameters. Therefore, the joint posterior distribution factors as

$$p(\theta_{0:n} | x_{1:n}) \propto p(x_{1:n} | \theta_{0:n}) p(b_{T,0:n}, \alpha_{T,0:n}) p(b_{P,0:n}, \alpha_{P,0:n})$$

$$\propto p(b_{T,0:n}, \alpha_{T,0:n} | x_{1:n}) p(b_{P,0:n}, \alpha_{P,0:n} | x_{P,1:n})$$

where (15) was used in the last step. This allows us to split the estimation problem into two independent estimation problems related to the P and T waves. In the following, only the T wave dynamic model and estimation problem are discussed, and the subscript T is omitted for notational convenience.

Due to the parametrization (14), the state parameter vector for the $n$th T wave interval is given by

$$\theta_n = \left( \begin{array}{l} b_n \\ \alpha_n \end{array} \right).$$

(Note that $\theta_n$ is now short for $\theta_{T,n}$, and thus different from the $\theta_n$ used, e.g., in Section 3.2.)

For the indicator vector $b_n$, we use the prior in (5) with $p_0 = p_1 = 1/(N_{T,n} + 1)$. This prior can be equivalently written as

$$\text{Pr} (b_n = \beta_k) = \frac{1}{N_{T,n} + 1}, \quad k \in \{0, \ldots, N_{T,n}\}$$

where $\beta_k$ for $k \in \{1, \ldots, N_{T,n}\}$ is an $N_{T,n} \times 1$ vector whose $k$th entry is 1 and all remaining entries are zero and $\beta_0$ is the all-zero vector (this represents the case where there is no T wave). Thus, the prior of $b_n$ is a uniform distribution on the set $\{\beta_0, \ldots, \beta_{N_{T,n}}\}$, which contains all possible $b_n$ such that $\|b_n\| = 1$ or $\|b_n\| = 0$. Indicator vectors $b_n$ for different time instants $n$ are assumed to be statistically independent.

Since the ECG waveforms are usually similar for two consecutive beats, we propose to assign a random walk prior to the T waveform coefficient vector $\alpha_n$, i.e.,

$$\alpha_n = \alpha_{n-1} + v_{n-1}$$

(16)

where $\alpha_{n-1}$ denotes the T waveform coefficient vector of the $(n - 1)$th beat and the vectors $v_n \sim \mathcal{N}(0, \sigma^2_{\alpha} I_G)$ are statistically independent (from each other and from $\alpha_{0:n}$) additive white Gaussian noise vectors. This leads to the prior

$$p(\alpha_n | \alpha_{n-1}) = \mathcal{N}(\alpha_{n-1}, \sigma^2_{\alpha} I_G)$$

(17)

which is the same as in the second case of (6). Note that here, in contrast to (6), the coefficient vector changes even if $\|b_n\| = 0$. The variance $\sigma^2_{\alpha}$ determines how fast the waveform coefficients
are expected to change with time. Since the non-QRS components are normalized by dividing by the amplitude of the respective R peak, we must account for possible significant variations of the waveforms with time. We therefore propose to use a large value of $\sigma^2_{\alpha}$, which corresponds to a non-informative conditional prior of $\alpha_n$. Note that the value of $\sigma^2_{\alpha}$ can be further adjusted either by an expert or by calculating the ECG waveform variance of an example ECG segment in an off-line parameter selection procedure as in [25]. Because of (16) and the independence of $v_n$ for different $n$ as well as from $\alpha_{0:n}$, the waveform coefficient vector $\alpha_n$ is conditionally independent, given $\alpha_{n-1}$, of all previous coefficient vectors $\alpha_{0:n-2}$, i.e.,

$$p(\alpha_n|\alpha_{0:n-1}) = p(\alpha_n|\alpha_{n-1}).$$

(18)

### 4.3 A marginalized particle filter for beat-to-beat wave analysis

Our goal is to estimate jointly the discrete-valued indicator vector $b_n$ and the waveform vector $\alpha_n$ i.e., to estimate the state vector $\theta_n$. In a Bayesian framework, all inference is based on the posterior distribution of the unknown parameters given the set of available observations, expressed as $p(\theta_{0:n}|x_{1:n})$. Particle filters (PFs) are a class of methods well-suited to perform the estimation of the hybrid state vector $\theta_{0:n}$. They approximate the target distribution by an empirical distribution

$$\hat{p}(\theta_{0:n}|x_{1:n}) = \sum_{i=1}^{N_n} w_n^{(i)} \delta \left( \theta_{0:n} - \theta_{0:n}^{(i)} \right), \quad \text{where} \quad \sum_{i=1}^{N_n} w_n^{(i)} = 1.$$ 

The weights $w_n^{(i)}$ and the particles $\theta_{0:n}^{(i)}$ are classically obtained by sequential importance sampling and a selection (resampling) step to prevent degeneracy [7].

While the classical PFs are fairly easy to implement, a main drawback is that, in practice, the required number of particles increases quickly with the state dimension. The MPF can reduce the number of parameters estimated by the PF and therefore allows fewer particles to be used. More specifically, the MPF takes advantage of linear Gaussian sub-structures in the state parameters $\theta_n$ to decrease the variance of the state estimates. The key idea is to split $\theta_n$ as

$$\theta_n = \begin{pmatrix} \theta_n^L \\ \theta_n^{NL} \end{pmatrix}$$

where $\theta_n^L$ denotes the state parameters with conditionally linear dynamics and $\theta_n^{NL}$ denotes the nonlinear state parameters. We can then marginalize out $\theta_n^L$ and generate particles distributed according to $p(\theta_n^{NL}|x_{1:n})$ using a PF. The particles are finally averaged to compute the MMSE estimator of $\theta_n^{NL}$. In parallel, each particle is associated with a Kalman filter (KF) that computes recursively the mean and covariance matrix of the Gaussian distribution $p(\theta_n^L|\theta_n^{NL}, x_{1:n})$.

In our case, it can be observed from (14) that both the discrete parameter vector $b_n$ and the continuous parameter vector $\alpha_n$ enter linearly in the observation $x_n$, given the respective other parameter. Since only continuous parameters can be handled by the KF, we choose $\theta_n^L = \alpha_n$ and $\theta_n^{NL} = b_n$. The KF and the PF correspond to two factors of the joint posterior according to the following factorization:

$$p(b_{0:n}, \alpha_{0:n}|x_{1:n}) = p(\alpha_{0:n}|b_{0:n}, x_{1:n}) p(b_{0:n}|x_{1:n}).$$

(19)

The marginal distribution of the discrete parameters is approximated by

$$\hat{p}(b_{0:n}|x_{1:n}) = \sum_{i=1}^{N_n} w_n^{(i)} \delta \left( b_{0:n} - b_{0:n}^{(i)} \right)$$

(20)
where \( N_s \) is the number of particles. Then, by inserting (20) in (19) and integrating out \( b_{0:n} \), the posterior distribution of the continuous parameters can be approximated by

\[
\hat{p}(\alpha_{0:n} | x_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} p(\alpha_{0:n} | b_{0:n}^{(i)}, x_{1:n}).
\]  

(21)

Integrating out \( \alpha_{0:n-1} \) yields

\[
\hat{p}(\alpha_n | x_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} p(\alpha_n | b_{0:n}^{(i)}, x_{1:n}).
\]  

(22)

It can be shown that \( p(\alpha_{0:n} | b_{0:n}^{(i)}, x_{1:n}) \) in (21) and \( p(\alpha_n | b_{0:n}^{(i)}, x_{1:n}) \) in (22) are Gaussian due to (14), (17), (18), the Gaussianity of the noise \( w_{n:k} \), and the relation

\[
p(\alpha_{0:n} | b_{0:n}, x_{1:n}) \propto p(x_{1:n} | \alpha_{0:n}, b_{0:n}) p(\alpha_{0:n} | b_{0:n}) = p(x_{1:n} | \alpha_{0:n}, b_{0:n}) p(\alpha_{0:n}).
\]

Therefore, (21) and (22) represent mixtures of Gaussian distributions. The mean of (22) can be expressed as

\[
\hat{E} \{ \alpha_n | x_{1:n} \} \triangleq \int \alpha_n \hat{p}(\alpha_n | x_{1:n}) \, d\alpha_n = \sum_{i=1}^{N_s} w_n^{(i)} \hat{\alpha}_n^{(i)},
\]  

(23)

with \( \hat{\alpha}_n^{(i)} = \hat{E} \{ \alpha_n | b_{0:n}^{(i)}, x_{1:n} \} = \int \alpha_n \hat{p}(\alpha_n | b_{0:n}^{(i)}, x_{1:n}) \, d\alpha_n \). This represents the proposed estimator of \( \alpha_n \); it is an approximation of the MMSE estimator. Note that one KF is associated with each particle \( b_{0:n}^{(i)} \) with \( i = 1, \ldots, N_s \). Furthermore, in practice, only the marginal distribution \( \hat{p}(\alpha_n | x_{1:n}) \) is updated (rather than \( \hat{p}(\alpha_{0:n} | x_{1:n}) \)). The MPF recursions are summarized in Algorithm 2. The different steps involved in this algorithm are detailed in the rest of this section.

### 4.3.1 Kalman filter prediction

At time \( n \), the previous MMSE state estimate is \( \hat{\alpha}_{n-1} = \hat{E} \{ \alpha_{n-1} | x_{1:n-1}, b_{0:n-1}^{(i)} \} \) and its covariance matrix is \( P_{n-1}^{(i)} = \text{Cov} \{ \alpha_{n-1} | x_{1:n-1}, b_{0:n-1}^{(i)} \} \). Define the predicted state vector \( \hat{\alpha}_{n|n-1}^{(i)} \triangleq \hat{E} \{ \alpha_n | x_{1:n-1}, b_{0:n-1}^{(i)} \} \) and its covariance \( P_{n|n-1}^{(i)} \triangleq \text{Cov} \{ \alpha_n | x_{1:n-1}, b_{0:n-1}^{(i)} \} \). Using (16), it can be shown that the prediction step of the KF can be written as

\[
\hat{\alpha}_{n|n-1}^{(i)} = \hat{\alpha}_{n-1}^{(i)}, \quad P_{n|n-1}^{(i)} = P_{n-1}^{(i)} + \sigma_n^2 I_G.
\]  

(24)

Note that the predicted state vector and its covariance computed by the KF, \( \hat{\alpha}_{n|n-1}^{(i)} \) and \( P_{n|n-1}^{(i)} \), will be directly used to propagate the particles and compute their importance weights, as explained presently (see (26)).

### 4.3.2 Importance distribution for the indicators

It is well known that the choice of the importance distribution is a critical issue in the design of efficient PF algorithms. To generate samples in relevant regions of the state space, i.e., corresponding to a high likelihood \( p(x_n | \theta_n) \), a natural strategy consists of taking into account information from the most recent observations \( x_n \). The importance distribution that is optimal in the sense that it minimizes the variance of the importance weights is \( q(b_n | b_{0:n-1}, x_{1:n}) = p(b_n | b_{0:n-1}, x_{1:n}) \) [9]. Thus, the optimal importance distribution for \( b_n \) is obtained as
Algorithm 2 Marginalized particle filter

{Initialization}

for \( i = 1, \ldots, N_s \) do

Set \( b_0^{(i)} = 0_{N_n \times 1}, P_0^{(i)} = 0_{G \times G} \), and \( w_0^{(i)} = 1 \), and choose a suitable initialization of the waveform coefficients \( \alpha_0^{(i)} \) (see Section 5.1.1).

end for

{Time recursion}

for \( n = 1, 2, \ldots \) do

for \( i = 1, \ldots, N_s \) do

{KF and PF propagation}

KF prediction for \( \alpha_n^{(i)} \) (see (24))

Sample \( b_n^{(i)} \sim \Pr (b_n = \beta_k | b_{0:n-1}^{(i)}, x_{1:n}^{(i)}) \) (see (25))

KF correction for \( \alpha_n^{(i)} \) (see (26))

Calculate weights

\[ \tilde{w}_n^{(i)} = w_n^{(i)} \sum_{k \in \mathcal{J}_{r,n}} \Pr (b_n = \beta_k | x_{1:n}^{(i)}) \Pr (b_n = \beta_k) \] (25)

end for

{Weight normalization}

for \( i = 1, \ldots, N_s \) do

\( w_n^{(i)} = \tilde{w}_n^{(i)} / \sum_{j=1}^{N_s} \tilde{w}_n^{(j)} \)

end for

{State estimation}

Estimation of \( b_n \) and \( \alpha_n \) (see (27))

{Particle resampling}

Calculate \( \hat{N}_{\text{eff}} = 1 / \sum_{i=1}^{N_s} \left( w_n^{(i)} \right)^2 \)

if \( \hat{N}_{\text{eff}} \leq 0.7 \cdot N_s \) then

Resample using systematic sampling scheme [7, p. 11]

end if

end for

\[
\Pr (b_n = \beta_k | b_{0:n-1}^{(i)}, x_{1:n}) \propto \Pr (x_n | b_n = \beta_k, b_{0:n-1}^{(i)}, x_{1:n-1}) \Pr (b_n = \beta_k).
\]

It can be shown that, for \( b_{0:n} \) given, \( \alpha_{0:n} \) and \( x_{1:n} \) are jointly Gaussian. It follows that the distribution \( \Pr (x_n | b_n = \beta_k, b_{0:n-1}^{(i)}, x_{1:n-1}) \) in (25) is a Gaussian one. According to (14), its mean \( \hat{x}_{n,k}^{(i)} \) and covariance matrix \( S_{n,k}^{(i)} \) can be computed from the KF prediction (24) as follows:

\[
\hat{x}_{n,k}^{(i)} = \tilde{B}_{n,k} H \hat{\alpha}_{n|n-1}^{(i)}
\]

\[
S_{n,k}^{(i)} = \tilde{B}_{n,k} H P_{n|n-1}^{(i)} H^T \tilde{B}_{n,k}^T + \sigma_w^2 I_{N_{T,n}}
\]

where \( \tilde{B}_{n,k} \) is the matrix \( \tilde{B}_n \), which corresponds to \( b_n = \beta_k \). Note that contrary to the standard PF, the importance distribution for the indicators no longer depends on the coefficient vector \( \alpha_{0:n} \), which has been marginalized out. On the other hand, it depends on the past sequence \( b_{0:n-1} \).
4.3.3 Kalman filter correction

After receiving the observation $x_n$ at time instant $n$, the waveform coefficients $\hat{\alpha}_n^{(i)}$ can be updated for each generated wave indicator particle $b_n^{(i)}$. The KF correction procedure can be written as

$$S_n^{(i)} = \tilde{B}_n^{(i)} H P_{n|n-1}^{n-1} H^T (\tilde{B}_n^{(i)})^T + \sigma_w^2 I_{N_{T,n}} \tag{26a}$$

$$K_n^{(i)} = P_{n|n-1}^{n-1} H^T (\tilde{B}_n^{(i)})^T S_n^{(i)}^{-1} \tag{26b}$$

$$\hat{\alpha}_n^{(i)} = \hat{\alpha}_{n|n-1}^{(i)} + K_n^{(i)} (x_n - \tilde{B}_n^{(i)} H \hat{\alpha}_{n|n-1}^{(i)}) \tag{26c}$$

$$P_n^{(i)} = (I_G - K_n^{(i)} \tilde{B}_n^{(i)} H) P_{n|n-1}^{n-1} \tag{26d}$$

where $\tilde{B}_n^{(i)}$ is the matrix $\tilde{B}_n$ which corresponds to $b_n = b_n^{(i)}$.

4.3.4 PF weight computation

When the optimal importance distribution is used to propagate the particles, the weights satisfy the following recursion:

$$w_n^{(i)} \propto w_{n-1}^{(i)} p(x_n|x_{1:n-1}, b_{0:n-1}^{(i)}) \tag{27}$$

Here, $p(x_n|x_{1:n-1}, b_{0:n-1}^{(i)})$ is the normalization constant of (25), i.e.,

$$p(x_n|x_{1:n-1}, b_{0:n-1}^{(i)}) = \sum_{k \in J_{T,n}} p(x_n|b_n^{(i)} = \beta_k, b_{0:n-1}^{(i)}, x_{1:n-1}) Pr(b_n^{(i)} = \beta_k).$$

4.3.5 State estimation

The sample-based blockwise MAP detector is used for estimating the binary sequence $b_n$, while the sample-based MMSE estimator (weighted sample mean, cf. (23)) is used for estimating the waveform coefficients $\alpha_n$:

$$\hat{b}_n = \arg\max_{i \in \{1, \ldots, N_s\}} \hat{p}(b_n^{(i)}|x_{1:n}), \quad \hat{\alpha}_n = \sum_{i=1}^{N_s} w_n^{(i)} \hat{\alpha}_n^{(i)}. \tag{27}$$

Here, $\hat{p}(b_n^{(i)}|x_{1:n})$ is obtained by marginalizing (20). The wave delineation consists of determining the peaks and boundaries of the detected P and T waves. As mentioned previously, the nonzero wave indicator estimated by the MPF directly indicates the center of the allotted waveform time window. Thus, the peak of the respective T or P wave is indicated by the location of the maximum of the estimated waveform. Furthermore, the wave boundaries can be located by applying the delineation criterion described in Section 3.4.3 to the estimated waveforms.

5 Simulation results

5.1 Simulation setup

Both of the proposed Bayesian wave detection/estimation/delineation methods were evaluated on the QT database (QTDB), which was previously used in several other studies [17].
QTDB provides a reference for wave-boundary validation. It is a two-channel database containing cardiologist annotations for at least 30 beats per dataset for both channels. It includes 105 datasets from the widely used MIT-BIH arrhythmia database, the European ST-T database, and some other well-known databases. The cardiologist annotations of the QTDB were performed using two leads, whereas the proposed delineation methods work on a single-channel basis. To compare the single-channel delineation results produced by our methods with the manual annotations of the QTDB, we chose for each T or P wave the channel where the detected peak location of the wave was closer to the annotated one (as suggested in [22, 2]). In a preprocessing step, the QRS complexes were detected and the borders of the non-QRS intervals \( J_n \) were determined using the Pan-Tompkins algorithm [23]. (The same preprocessing step was performed in [20, 21].) In another preprocessing step, baseline wanderings were removed. P and T search intervals \( J_{T,n} \) and \( J_{P,n} \) were then defined as the first and second half of \( J_n \). Both of the proposed methods sequentially process one non-QRS interval \( J_n \) after another.

### 5.1.1 BGS setup

For each non-QRS interval, the BGS generated 100 samples according to the conditional distributions specified in Section 3.3.2. The first 40 samples constituted the burn-in period, and the remaining 60 were used for detection/estimation (thus, \( N_s = 60 \)). The fixed hyperparameters involved in the prior distributions were chosen as \( p_0 = 0.01, \sigma^2_\alpha = 0.01, \xi = 11, \) and \( \eta = 0.5 \); these values allow for an appropriate waveform variability from one beat to another and provide a noninformative prior for the noise variance \( \sigma^2_w \). Note that the non-QRS components are normalized using the corresponding R peak values to handle different amplitude resolutions. For the first non-QRS interval (\( n = 1 \)), the previous waveform coefficient estimates \( \hat{\alpha}_{T,n-1} \) and \( \hat{\alpha}_{P,n-1} \) were initialized with the coefficient vector \( \alpha \) for which \( h \) is closest to the \( 2L + 1 \) Hann window [12], with an amplitude equal to half the R peak amplitude. The waveform length is chosen as \( 2L + 1 = N_n / 3 \), which is large enough to accommodate the actual support of the T or P wave. Because the proposed beat-to-beat BGS method processes only one non-QRS interval at any given time, both its memory requirements and its computational complexity are smaller than those of the window-based method of [20]. For instance, for the proposed method using 100 sampler iterations, the processing time per beat is approximately 0.3s using a nonoptimized MATLAB implementation running on a 3.0-GHz Pentium IV computer, compared to about 2s for the method of [20].

### 5.1.2 MPF setup

In the MPF method, the fixed hyperparameters involved in the prior distributions were chosen as \( \sigma^2_\alpha = 0.01 \) and \( \sigma^2_w = 0.1 \). This value of \( \sigma^2_\alpha \) allows for an appropriate waveform variability from one beat to another. The chosen value of \( \sigma^2_w \) comes from a previous estimation of the noise level, e.g., using the mean value estimated by the BGS method, but could be taken from any other noise estimator. Recall that the MPF method assumes \( \sigma^2_w \) to be known, whereas the BGS method does not. The non-QRS components are again normalized using the corresponding R peak values to handle different amplitude resolutions. The waveform vector \( h_0 = H\hat{\alpha}_0 \) was initialized as in the BGS method, i.e., with a length \( 2L + 1 = N_n / 3 \) Hann window whose amplitude was half the R peak amplitude. An important issue with PF methods is the number of particles. Table 1 shows the normalized mean square error (NMSE) of the estimated non-QRS components versus the number of particles \( N_s \). As can be seen, benefiting from the optimal importance distribution mentioned in Section 4.3.2, good estimation performance can be obtained with a moderate number of particles. For the MPF method using 200 particles, the processing time per beat is approximately 0.5s using a nonoptimized MATLAB implementation.
running on a 3.0-GHz Pentium IV computer. We have chosen \( N_s = 200 \) particles for all the following simulations in order to guarantee an NMSE close to -40dB.

Table 1: Normalized mean square error (NMSE) versus number of particles used in the MPF method.

<table>
<thead>
<tr>
<th>( N_s )</th>
<th>10</th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMSE</td>
<td>-25dB</td>
<td>-31dB</td>
<td>-34dB</td>
<td>-40dB</td>
<td>-42dB</td>
</tr>
</tbody>
</table>

5.2 Qualitative analysis

In this section, we first show the posterior distributions as well as estimation and delineation results obtained by the proposed beat-to-beat BGS method on a typical example. Then, we present a qualitative comparison of the proposed BGS and MPF methods with state-of-the-art methods on several representative ECG segments.

5.2.1 Example of posterior distributions, estimation and delineation

Fig. 3(a) shows two consecutive heartbeats from the QTDB dataset sele0136.

Figure 3: (a) Two consecutive heartbeats from QTDB dataset sele0136; (b) estimated marginal posteriors \( P_S(b_{T,n,k}=1) \) obtained by the proposed BGS method; (c) estimated P and T waveforms and delineation results obtained by the proposed BGS method.

The corresponding sample-based estimates of the marginal posterior probabilities of having a T or P wave at a given location \( k \), \( P_S(b_{T,n,k}=1) \) and \( P_S(b_{P,n,k}=1) \), are depicted in Fig. 3(b). (For \( k \in J_T \), \( P_S(b_{k}=1) \) equals the probability \( P_S(b_{T}) \) of the specific hypothesis \( b_{T} \) that contains a 1-entry at location \( k \), and similarly for \( k \in J_P \).) Fig. 3(c) shows the P and T waveforms estimated by the proposed BGS method for each search interval along with the corresponding delineation results (i.e., the estimated wave onsets, peaks, and ends, which were determined...
as described in Section 3.4.3). As can be seen, there are noticeable differences between the two consecutive T waveform estimates (at time instants 4.92s and 6.10s), as well as between the two consecutive P waveform estimates (at time instants 5.64s and 6.83s). This confirms the pseudo-stationary nature of the ECG signal and justifies our introduction of a beat-to-beat processing scheme that allows for beat-to-beat variations of the P and T waveforms.

5.2.2 Benefits of the beat-to-beat processing

Then, we present a qualitative comparison of the proposed BGS and MPF methods with the multi-beat method of [20] (based on a partially collapsed Gibbs sampler (PCGS)) to highlight the benefits of beat-to-beat processing. To evaluate the methods under real physiological noise conditions, we added muscular activity noise from the MIT-BIH noise stress test database. The estimated non-QRS signal components obtained with the different methods are displayed in Figs. 4 and 5 for eight successive beats of a segment of QTDB dataset sele0136.

![Figure 4](image)

Figure 4: Four consecutive segments from QTDB dataset sele0136 (blue) superimposed with their estimates (dotted red). (a) PCGS multi-beat method of [10]; (b) proposed beat-to-beat BGS method; (c) proposed beat-to-beat MPF method.

The original ECG signal is also shown for comparison. It can be seen that the proposed beat-to-beat methods (BGS and MPF) provide closer agreements with the original ECG signal when compared to the multi-beat method, especially at the onsets and ends of the waves, which is a desirable property for wave delineation. These results show that, contrary to the multi-beat method of [20], the beat-to-beat BGS and MPF methods are able to capture the changes affecting the P and T waveforms. Whereas the delineation results (locating the peaks and estimating the boundaries of each wave) obtained with the two proposed methods are similar,
Figure 5: Four other consecutive segments from QTDB dataset sele0136 (blue) superimposed with their estimates (dotted red). (a) PCGS multi-beat method of [10]; (b) proposed beat-to-beat BGS method; (c) proposed beat-to-beat MPF method.
the MPF outperforms the BGS from a waveform estimation point of view, at the price of a higher computational cost.

5.2.3 Comparison with state-of-the-art methods

Next, we present a qualitative comparison of the proposed BGS and MPF methods with the Gaussian mixture model and extended Kalman filter (EKF) method of [26] and the multi-beat method of (based on a partially collapsed Gibbs sampler) [20]. To evaluate the two methods under real physiological noise conditions, we have added muscular activity noise from the MIT-BIH noise stress test database. Fig. 6(a) shows a segment of QTDB dataset sele0136. Fig. 6(b) shows the same segment corrupted by muscular activity noise with a signal-to-noise ratio (SNR) of 10dB. Fig. 6(c), 6(d), 6(e), and 6(f) show the estimated non-QRS signal component obtained from the noisy signal by the EKF method of [26], the multi-beat method of [20], the proposed beat-to-beat BGS method, and the proposed MPF method, respectively. The original ECG signal is also shown in Fig. 6(c)–(f) for comparison. It can be seen that the proposed beat-to-beat methods (BGS and MPF) provide closer agreements with the original ECG signal when compared to the other methods, especially at the onsets and ends of the waves, which is a desirable property for wave delineation.

5.2.4 Example of a pathological ECG

Fig. 7 and Fig. 8 show analogous results for a segment of QTDB dataset sel803 that contains premature ventricular contractions (a pathology which has parts of the T waves crossing the interval border and the P waves missing). The proposed methods are seen to exhibit good performance even in the presence of nonmonotonic morphological abnormalities. For the beat-to-beat BGS method, since only the locations of the wave peaks are constrained to lie within their respective search intervals whereas the non-QRS signal component is processed jointly, this border situation can be properly handled. For the MPF method, this situation is handled by including in the observations $\tau_T$ (associated with the T-wave interval) the first $L$ samples of the following P wave interval, where $L$ is half of the allotted waveform length. More precisely, the first $L$ samples of $J_P$ are considered to complete the end of $J_T$. This allows a complete T waveform estimation even if the wave indicator is located at the last position within $J_T$. Since one non-QRS interval is processed sequentially from a T wave interval to a P wave interval, the estimated T wave portion within the overlapped part can be extracted. Again, it can be seen that the estimates obtained with the two proposed methods are quite similar, with slightly better estimates produced by the MPF (as demonstrated in Fig. 8), and they are slightly better than those obtained with the other two methods.

5.3 Quantitative analysis

Next, we provide a quantitative performance comparison of the two proposed methods with the multi-beat method of [20] and three alternative methods [16, 22, 32], based on an exhaustive evaluation performed on the entire QTDB. For a quantitative analysis of the performance of P and T wave detection, as in [16, 22, 32, 20, 21], we have computed the sensitivity (also referred to as detection rate) $Se = TP/(TP + FN)$ and the positive predictivity $P^+ = TP/(TP + FP)$, where TP denotes the number of true positive detections (wave was present and was detected), FN stands for the number of false negative detections (wave was present but was missed), and FP for the number of false positive detections (wave was not present but was detected). The performance of wave delineation has been measured by the average (denoted as $m$) and standard deviation (denoted as $s$) of the time differences between the results of the considered method.
Figure 6: (a) Segment from QTDB dataset sele0136; (b) noisy version including muscular activity noise with SNR = 10dB; (c) non-QRS signal component estimated by the EKF method of [26] (red) and original signal (blue); (d) non-QRS signal component estimated by the multi-beat method of [20] (red) and original signal (blue); (e) non-QRS signal component estimated by the proposed beat-to-beat BGS method (red) and original signal (blue); (f) non-QRS signal component estimated by the proposed beat-to-beat MPF method (red) and original signal (blue).
Figure 7: (a) Segment from QTDB dataset sel803; (b) noisy version including muscular activity noise with SNR = 10dB; (c) non-QRS signal component estimated by the EKF method of [26] (red) and original signal (blue); (d) non-QRS signal component estimated by the multi-beat method of [20] (red) and original signal (blue); (e) non-QRS signal component estimated by the proposed beat-to-beat BGS method (red) and original signal (blue); (f) non-QRS signal component estimated by the proposed beat-to-beat MPF method (red) and original signal (blue).
estimated non-QRS intervals by using beat-to-beat block Gibbs sampler

estimated non-QRS intervals by using the beat-to-beat particle filter

Figure 8: Segment from QTDB dataset sel803 (detail from Fig. 7): (a) non-QRS signal component estimated by the proposed beat-to-beat BGS method (red) and original signal (blue); (b) non-QRS signal component estimated by the proposed beat-to-beat MPF method (red) and original signal (blue).

and the corresponding cardiologist annotations. The indicated time values (in ms) are based on a sampling frequency of 250Hz. The quantities \( m \) and \( s \) were computed separately for the wave onset times \( t_{P,\text{on}} \) and \( t_{T,\text{on}} \), the wave peak times \( t_{P,\text{peak}} \) and \( t_{T,\text{peak}} \), and the wave end times \( t_{P,\text{end}} \) and \( t_{T,\text{end}} \). We note that while the QTDB includes annotations made by two cardiologists, we have considered only those of the first cardiologist, who provided annotations for at least 30 beats per dataset.

Table 2 shows the results for \( Se, P^+ \), and \( m \pm s \) obtained for the entire QTDB. It can be seen that the two proposed methods detect the P and T waves annotated by the cardiologist with high sensitivity: the sensitivity \( Se \) is 100% for the T waves and 99.93% or 99.95% for the P waves. Similarly good results were obtained for the positive predictivity \( P^+ \), which is between 98.01% and 99.30% for the T waves and 99.10% or 99.23% for the P waves. Both the Se values and the \( P^+ \) values are typically better than those obtained with the other methods, including the recently proposed multi-beat method of [20].

Regarding the delineation performance, it is seen from Table 2 that the two proposed methods delineate the annotated P and T waves with mean errors \( m \) not exceeding 4ms (except for \( t_{T,\text{on}} \)) and with smaller standard deviations \( s \) than the other methods (with two exceptions). We note that delineation error tolerances have been recommended by the CSE Working Party [30]. In particular, the standard deviation \( s \) for \( t_{P,\text{on}}, t_{P,\text{end}}, \) and \( t_{T,\text{end}} \) should be at most \( 2s_{\text{CSE}} \), which is listed in the last row of Table 2. However, a stricter recommendation proposed in [22] is \( s \leq s_{\text{CSE}} \). According to Table 2, the standard deviations for \( t_{P,\text{end}} \) achieved by both proposed methods and the standard deviation for \( t_{P,\text{on}} \) achieved by the proposed MPF method comply with the loose recommendation. For the \( t_{T,\text{end}} \) results, both proposed methods comply with the
Table 2: Comparison of the detection and delineation performance of the proposed beat-to-beat BGS and MPF methods with the multi-beat method of [20], the wavelet transform based method of [22] (WT), the low-pass differentiation based method of [16] (LPD), and the action potential based method of [32]. The variances of these methods are compared with the delineation error tolerance of [30], which is provided in the last row. (N/A: not available)

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameters</th>
<th>$t_{P,\text{on}}$</th>
<th>$t_{P,\text{peak}}$</th>
<th>$t_{P,\text{end}}$</th>
<th>$t_{T,\text{on}}$</th>
<th>$t_{T,\text{peak}}$</th>
<th>$t_{T,\text{end}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beat-to-beat annotations</td>
<td>3176</td>
<td>3176</td>
<td>3176</td>
<td>1345</td>
<td>3403</td>
<td>3403</td>
<td></td>
</tr>
<tr>
<td>BGS (proposed)</td>
<td>Se (%)</td>
<td>99.93</td>
<td>99.93</td>
<td>99.93</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>99.10</td>
<td>99.10</td>
<td>99.10</td>
<td>98.01</td>
<td>99.30</td>
<td>99.30</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>3.4±14.2</td>
<td>1.1±5.3</td>
<td>-2.1±9.8</td>
<td>6.8±16.3</td>
<td>-0.8±4.1</td>
<td>-3.1±14.0</td>
</tr>
<tr>
<td>Beat-to-beat annotations</td>
<td>3176</td>
<td>3176</td>
<td>3176</td>
<td>1345</td>
<td>3403</td>
<td>3403</td>
<td></td>
</tr>
<tr>
<td>MPF (proposed)</td>
<td>Se (%)</td>
<td>99.95</td>
<td>99.95</td>
<td>99.95</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>99.23</td>
<td>99.23</td>
<td>99.23</td>
<td>98.67</td>
<td>99.20</td>
<td>99.20</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>3.1±8.3</td>
<td>1.2±5.3</td>
<td>2.7±9.8</td>
<td>6.5±16.3</td>
<td>-0.4±4.8</td>
<td>-3.8±14.2</td>
</tr>
<tr>
<td>Multi-beat annotations</td>
<td>3176</td>
<td>3176</td>
<td>3176</td>
<td>1345</td>
<td>3403</td>
<td>3403</td>
<td></td>
</tr>
<tr>
<td>partially collapsed Gibbs sampler [20]</td>
<td>Se (%)</td>
<td>99.60</td>
<td>99.60</td>
<td>99.60</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>98.04</td>
<td>98.04</td>
<td>98.04</td>
<td>97.23</td>
<td>99.15</td>
<td>99.15</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>1.7±10.8</td>
<td>2.7±8.1</td>
<td>2.5±11.2</td>
<td>5.7±16.5</td>
<td>0.7±9.6</td>
<td>2.7±13.5</td>
</tr>
<tr>
<td>WT [22]</td>
<td>annotations</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Se (%)</td>
<td>98.87</td>
<td>98.87</td>
<td>98.75</td>
<td>99.77</td>
<td>99.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>91.03</td>
<td>91.03</td>
<td>91.03</td>
<td>N/A</td>
<td>97.79</td>
<td>97.79</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>2.0±14.8</td>
<td>3.6±13.2</td>
<td>1.9±12.8</td>
<td>N/A</td>
<td>0.2±13.9</td>
<td>-1.6±18.1</td>
</tr>
<tr>
<td>LPD [16]</td>
<td>annotations</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Se (%)</td>
<td>97.70</td>
<td>97.70</td>
<td>97.70</td>
<td>99.00</td>
<td>99.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>91.17</td>
<td>91.17</td>
<td>91.17</td>
<td>N/A</td>
<td>97.74</td>
<td>97.74</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>14.0±13.3</td>
<td>4.8±10.6</td>
<td>-0.1±12.3</td>
<td>N/A</td>
<td>-7.2±14.3</td>
<td>13.5±27.0</td>
</tr>
<tr>
<td>Action potential based method</td>
<td>annotations</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>[32]</td>
<td>Se (%)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>92.60</td>
<td>92.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P^+ %$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>$m \pm s \text{ (ms)}$</td>
<td>N/A</td>
<td>N/A</td>
<td>20.9±29.6</td>
<td>-12.0±23.4</td>
<td>0.8±30.3</td>
<td></td>
</tr>
<tr>
<td>Delineation error tolerance</td>
<td>$2\sigma_{\text{CSE}} \text{ (ms)}$</td>
<td>10.2</td>
<td>N/A</td>
<td>12.7</td>
<td>N/A</td>
<td>N/A</td>
<td>30.6</td>
</tr>
</tbody>
</table>

strict recommendation.

From Table 2, it is furthermore seen that the detection and delineation results obtained with the two proposed methods are quite similar. The computational complexity of the two methods is comparable but slightly higher for the MPF. Despite the moderate number of particles needed by the MPF, the fact that one KF is associated with each particle introduces matrix inversion operations which make the MPF more complex than the BGS. For the MPF using 200 particles, the processing time per beat is approximately 0.5s using a non-optimized MATLAB implementation running on a 3.0-GHz Pentium IV computer, compared to about 0.3s for the BGS.

6 Conclusion

This report presented and studied two Bayesian methods for beat-to-beat P and T wave delineation and waveform estimation. We introduced two Bayesian models for the non-QRS components of the ECG signal. Instead of using a processing window that contains several successive beats involving the same P and T waveforms, the proposed methods account for beat-to-beat
variations of the P and T waveforms by processing individual beats sequentially (i.e., with memory). First, a block Gibbs sampler method was proposed to estimate the unknown parameters of the beat-to-beat Bayesian model. Alternatively, in order to take advantage of all the available information contained in the past of the beat to be processed, a dynamic model was proposed. This model exploits the sequential nature of the ECG signal by using a random walk model for the waveform coefficients. A marginalized particle filter was then proposed to estimate the unknown parameters of the dynamic model.

The main features and contributions of this work can be summarized as follows:

1. Beat-to-beat block Gibbs sampler (BGS) method

   - The proposed Bayesian model uses the P and T waveform estimates of the previous beat as prior information for detecting/estimating the current P and T waves.
   - By properly accounting for the strong local dependencies in, and the sequential nature of, ECG signals, the proposed BGS exhibits a faster convergence than the samplers used in [20, 21].
   - The high accuracy of the proposed technique for P and T waveform estimation allows a threshold-free delineation technique to be used.
   - The beat-to-beat processing mode leads to smaller memory requirements and a lower computational complexity compared to the window-based Bayesian methods in [20, 21].

2. Marginalized particle filter (MPF) method

   - The sequential nature of the ECG signal is exploited by using a dynamic model within the Bayesian framework.
   - The proposed MPF method efficiently estimates the unknown parameters of the dynamic model. Thanks to the marginalization, a smaller number of particles is needed for good estimation performance, compared to the classical particle filter.
   - Compared to the beat-to-beat BGS method, the MPF method is potentially advantageous in that it considers all the available beats in the waveform estimation.

The proposed beat-to-beat Bayesian methods were validated using the QT database. A comparison with the multi-beat method of [20] and with other benchmark methods demonstrated that both proposed methods can provide significant improvements regarding P and T wave detection rate, positive predictivity, and delineation accuracy. We note that the proposed methods are single-lead based ECG processing methods. They can be extended to multi-lead ECG signals by including post-processing decision rules to determine global marks from the single-lead delineation results [1].

Besides its suitability for real-time ECG monitoring, another advantage of the proposed beat-to-beat processing mode is the possibility of analyzing the beat-to-beat variation and evolution of the P and T waveforms. Potential clinic applications include T wave alternans (TWA) detection in intra-cardiac electrograms. This application is currently under investigation.

A Derivation of BGS sampling distributions

This appendix derives the expressions of two of the sampling distributions provided in Section 3.3.2. We recall the definitions \( \theta \triangleq (b_T^T, b_T^a, \alpha_T^T, \alpha_T^a, \sigma_w^2)^T \) and \( \bar{x}_T \triangleq x - B_T H \alpha_P \). As before, the subscript \( n \) is suppressed.
A.1 Wave indicators

To derive expression (9) of the sampling distribution for \( b_T \), we first note that

\[
p(b_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) = \int p(b_T | \alpha_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \, d\alpha_T
\]

\[
= \int p(b_T | \alpha_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \, d\alpha_T
\]

\[
\propto \int p(\theta | x, \alpha_{T,n-1}, \alpha_P, n-1) \, d\alpha_T
\]

\[
\propto \int p(x | \theta) p(\theta | \alpha_{T,n-1}, \alpha_P, n-1) \, d\alpha_T.
\]

In the second last and last steps, we used Bayes’ rule and (8), respectively. Next, we insert (4) and (7) and drop factors that do not depend on \( b_T \) or \( \alpha_T \). This yields

\[
p(b_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \int \exp \left( -\frac{\| \hat{x}_T - B_T H \alpha_T \|^2}{2\sigma_w^2} \right) p(\alpha_T | b_T, \alpha_{T,n-1}) \, d\alpha_T.
\]  

(28)

Now, we insert (6) for \( p(\alpha_T | b_T, \alpha_{T,n-1}) \) and (5) for \( p(b_T) \). For the case where \( \| b_T \| = 0 \) (no T wave in \( f_T \)), this yields

\[
p(b_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} \right) \left[ \int \delta(\alpha_T - \hat{\alpha}_{T,n-1}) \, d\alpha_T \right] p_0 = \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} \right) p_0.
\]  

(29)

For the case where \( \| b_T \| = 1 \) (one T wave in \( f_T \)), we obtain

\[
p(b_T | b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \frac{1}{(\sqrt{2\pi\sigma_\alpha})^2} \left[ \int \exp \left( -\frac{\| \hat{x}_T - B_T H \alpha_T \|^2}{2\sigma_w^2} \right) \, d\alpha_T \right] p_1
\]

\[
= \frac{1}{(\sqrt{2\pi\sigma_\alpha})^2} \left[ \int \exp \left( -\frac{1}{2} (\alpha_T - \mu_1^T \Sigma_1^{-1} (\alpha_T - \mu_1)) \right) \, d\alpha_T \right] \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} + \mu_1 \Sigma_1^{-1} \mu_1 \right) p_1
\]

\[
= \sqrt{\frac{\Sigma_1}{\sigma_\alpha}^2} \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} \right) \exp(\mu_1 \Sigma_1^{-1} \mu_1) p_1
\]  

(30)

with \( \mu_1 \) and \( \Sigma_1 \) as defined in (10) and (11), respectively. In the third case, \( \| b_T \| > 1 \), the prior in (5) is zero, hence (28) is also zero. Since the constant factor \( \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} \right) \) appears in both (29) and (30), we obtain (9).

A.2 Waveform coefficients

To derive expression (12) of the sampling distribution for \( \alpha_T \), we again use Bayes’ rule as well as equations (8) and (7). We thus obtain

\[
p(\alpha_T | b_T, b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto p(\theta | x, \alpha_{T,n-1}, \alpha_P, n-1) \propto p(x | \theta) p(\theta | \alpha_{T,n-1}, \alpha_P, n-1) \propto p(x | \theta) p(\alpha_T | b_T, \alpha_{T,n-1}).
\]

In the last expression, we dropped all factors that do not depend on \( \alpha_T \). Inserting (4) for \( p(x | \theta) \) and (6) for \( p(\alpha_T | b_T, \alpha_{T,n-1}) \), we obtain for the case \( \| b_T \| = 0 \)

\[
p(\alpha_T | b_T, b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \exp \left( -\frac{\| \hat{x}_T \|^2}{2\sigma_w^2} \right) \delta(\alpha_T - \hat{\alpha}_{T,n-1}) \propto \delta(\alpha_T - \hat{\alpha}_{T,n-1})
\]

and for the case \( \| b_T \| = 1 \)

\[
p(\alpha_T | b_T, b_p, \alpha_{T,n-1}, \alpha_P, \sigma_w^2, x) \propto \frac{1}{(\sqrt{2\pi\sigma_\alpha})^2} \exp \left( -\frac{\| \hat{x}_T - B_T H \alpha_T \|^2}{2\sigma_w^2} \right) \propto N(\mu_1, \Sigma_1).
\]  

26
References


